(51) 国際特許分類7 C12N 15/12, C07K 14/47, 16/18	A1	(11) 国際公開番号	WO00/29572
		(43) 国際公開日	2000年5月25日(25.05.00)

(21) 国際出願番号　　　　　　PCT/JP99/06388 (22) 国際出願日　　　　1999年11月16日(16.11.99) (30) 優先権データ 特願平10/325658　1998年11月16日(16.11.98)　JP 特願平10/325659　1998年11月16日(16.11.98)　JP 特願平11/301582　1999年10月22日(22.10.99)　JP (71) 出願人　（米国を除くすべての指定国について） 科学技術振興事業団 (JAPAN SCIENCE AND TECHNOLOGY CORPORATION) [JP/JP] 〒332-0012 埼玉県川口市本町4丁目1番8号 Saitama, (JP) (72) 発明者；および (75) 発明者／出願人　（米国についてのみ） 髙井義美(TAKAI, Yoshimi)[JP/JP] 〒651-2102 兵庫県神戸市西区学園東町2丁目5番地の73 Hyogo, (JP) 畑　　裕(HATA, Yutaka)[JP/JP] 〒651-2112 兵庫県神戸市西区大津和1丁目1番地の8 ハイツ大津和609 Hyogo, (JP)	矢尾育子(YAO, Ikuko)[JP/JP] 〒651-2111 兵庫県神戸市西区池上2丁目25の8 グレースハウスムーン206 Hyogo, (JP) (74) 代理人 弁理士　西澤利夫(NISHIZAWA, Toshio) 〒150-0042 東京都渋谷区宇田川町37-10 麻仁ビル6階 Tokyo, (JP) (81) 指定国　　CA, US 添付公開書類 　　国際調査報告書

(54)Title:　　PROTEINS MAGUIN

(54)発明の名称　　蛋白質MAGUIN

(57) Abstract

　　A rat protein MAGUIN-1 having the amino acid sequence represented by SEQ ID NO:1 and another rat protein MAGUIN-2 having the amino acid sequence represented by SEQ ID NO:3 which are unknown molecules capable of binding specifically to S-SCAM playing an important role and participating in the postsynaptic PSD formation; rat genes encoding these proteins; cDNAs of these proteins which are DNA fragments respectively having the base sequences represented by SEQ ID NOS:2 and 4; recombinant vectors carrying these DNA fragments respectively; and an antibody against the above protein MAGUIN-1.

DETAILED DESCRIPTION

[Detailed Description of the Invention]
[0001]
[Field of the Invention] This invention relates to protein MAGUIN-1 specifically combined with component PSD-95/SAP90 and S-SCAM of a nerve synaptic connection posterior part, and MAGUIN-2 which are that isoform. This invention relates to the elucidation of mechanisms, such as formation of a human nervous system, development or storage, and study, and the functional or new protein of a nervous system useful to the ED for a diagnosis of the various nervous diseases which consider a structural failure as a cause, prevention, and a therapy etc. further in more detail.
[0002]
[Description of the Prior Art] system control is very important for normal functional maintenance of multicellular organisms including Homo sapiens at the time of a series of life processes, such as formation of the cell junction supporting information interchange between the cells which follow orderly cell adhesion and orderly it, and differentiation of the cell caused based on the information which delivers through the cell junction further and is carried out.
[0003] Although much similarity exists in the cell junction accepted in a multicellular organism, the singularity by the class of cell also exists by one side, and especially nerve synaptic connection has the greatly different description from other cell junction. In the first place, in the first place, nerve synaptic connection has a polarity, and structures differ remarkably in the anterior part and posterior part. The back synapse film has the unique cytoskeleton structure called postsynaptic density (PSD) to the presynapsis film having the structure in connection with emission of neurotransmitter (Curr.Opin.Neurobiol.3, 732-737, 1993; Trends Neurosci.20, 264-268, 1997). Also once nerve synaptic connection is materialized, I hear that it shows change (plasticity) of the description of association to the second in connection with neurotransmission, and there is in it. It is thought that this plasticity is the phenomenon of accomplishing the base of the mechanism of storage and study, and it is the description which just makes nerve synaptic connection conspicuous from other cell junction. Although the detail of a reversible mechanism is not yet clear, existence of the presynapsis plasticity depending on the anterior part of nerve synaptic connection and the back synaptic plasticity depending on a posterior part is known. Anyway, it is predicted that the characteristic structure of nerve synaptic connection is involving deeply, and the elucidation of this structure is considered to be indispensable by the reversible mechanism and the pan at the mechanism elucidation of storage and study.
[0004] For effective neural transmission, it is required for a neurotransmitter receptor to exist in an exact location in the back synapse film. From the latest research, some molecules which function on are recording and clustering of various acceptors in PSD are identified. For example, PSD-95/SAP90 and its isoform (isoforms) are combined with an NMDA receptor and a Shaker mold K+ channel (for example, Neuron 9, 929-942, 1993), and what (Nature 386, 279-284, 1997; Nature 386, 284-288, 1997) GRIP and Homer are combined with an AMPA acceptor and a metabolic glutamate receptor for, respectively is known. The common description of these molecules is that they have a module PDZ domain for a protein-protein interaction (for example, TrendsBiochem.Sci.20, 102-103, 1995; Trends Biochem.Sci.20, 350, 1995; Curr.Biol.6, 1385-1388, 1996; Neuron 17, 575-578, 1996).
[0005] A PDZ domain is found out as a repeat of about 90 amino acid of PSD-95/SAP90 at first. Then, three protein [PSD-95/SAP90 and discs-large of drosophila (Drosophilia) Neoplasm control protein Dlg-A, Since it was contained in tight junction protein ZO-1, it was named PDZ (). [Neuron 9, 929-942, 1992;J.Biol.Chem.268, 4580-4583,] [1993;] Cell 66, 451-464, 1991; J.Cell Biol.121, 491-502 ; Proc.Natl.Acad.Sci.USA.90, 7834-7838, 1993. This PDZ domain is combined with the C terminal array of various protein. That is, the 1st of PSD-95/SAP90 and the 2nd PDZ domain are combined with an NMDA receptor and a Shaker mold K+ channel (Science 269, 1737-1740, 1995; Nature 378, 85-88, 1995). The

3rd PDZ domain is combined with the cytoplasm domain of the nerve cell adhesion molecules called neuro-RIGIN (neuroligin) (Science 277, 1511-1515, 1997). Moreover, PSD-95/SAP90 forms a gay tetramer through DISARUFIDO association of an amino terminal (Neuron 18, 803-814, 1997).

[0006] PSD-95/SAP90 has one SH3 domain and one guanylate kinase (GK) domain other than a PDZ domain. Although SH3 domain was specified as a signal transfer module of src tyrosine kinase, it is identified also in much protein besides after that (Cell 80, 237-248, 1995). However, the function of SH3 domain in PSD-95/SAP90 is unknown. GK domains are yeast guanylate kinase and homologous, and exist in various film related protein. Such protein is considered to maintain specific membrane structure, and is named film related guanylate kinase (membrane-associated guanylate kinase:MAGUK) (Mech.Dev.44, 85-89, 1993; Curr.Biol.6, 382-384, 1996).

[0007] In recent years, the molecule which interacts with GK domain of PSD-95/SAP90 is identified by three research consortia, and is named GKAP/SAPAP/DAP (J. Cell Biol.136, 669-678, 1997; J.Biol.Chem.272, 11943-11951, 1997; Genes Cells 2, 415-424, 1997; J.Neurosci.17, 5687-5696, 1997). Among these molecules, SAPAP (1 2) is the molecule which the artificers of this application found out, and has already carried out patent application (Japanese Patent Application No. No. 11714 [nine to], Japanese Patent Application No. No. 11715 [nine to]). Moreover, in HEK293 cell, this SAPAP combines with a cell membrane firmly, and even if the artificers of this application find out making PSD-95/SAP90 shift to a cell membrane, there are (J. Biol. Chem.272, 11943-11951, 1997).

[0008] And the artificers of this application have found out further the molecule specifically combined with GKAP/SAPAP/DAP which participates in PSD formation of a nerve synaptic connection posterior part. It has the structure of MAGUK, has GK domain of an amino terminal, two WW domains, and six PDZ domains, and combines with GKAP/SAPAP/DAP through GK domain, and these GKAP/SAPAP/DAP tie molecules are combined with an NMDA receptor and neuro-RIGIN (neuroligin) through a PDZ domain. And since the acceptor and cell adhesion protein of synaptic connection are unified, it is named S-SCAM (synaptic scaffolding molecule), and patent application of this protein has already been carried out (Japanese Patent Application No. No. 302239 [ten to]).

[0009]
[Problem(s) to be Solved by the Invention] It is considered by various knowledge submitted by the artificers of this application, and other research consortia for GKAP/SAPAP/DAP to be the nucleus-protein of PSD, and it is indispensable to specify the new molecule which interacts with S-SCAM, in order to solve the whole aspect of PSD formation, although it is shown clearly that S-SCAM combined with this GKAP/SAPAP/DAP has played the role important for PSD formation.

[0010] Moreover, it is expected that about [that identification and the elucidation of a function of such a molecule bring a big advance to an understanding of higher brain functions, such as storage and study,], this molecule or the activity of this molecule, various matter that influences to an operation, a compound, etc. contribute functionally [central nervous system] greatly [list / various cause elucidations of a nervous disease which consider a structural failure as a cause, and / development / of a diagnosis of these diseases, prevention, a cure, etc.]. [Homo sapiens]

[0011] Invention of this application is made in view of the situation as above, and aims at offering the strange molecule specifically combined with S-SCAM which has played the important role which participates in PSD formation of a nerve synaptic connection posterior part.

[0012]
[Means for Solving the Problem] This application offers invention of the following (1) - (9) as what solves the above-mentioned technical problem.

(1) Rat protein MAGUIN-1 which has the amino acid sequence of the array number 1.

(2) Rat protein MAGUIN-2 which have the amino acid sequence of the array number 3.

(3) Rat protein MAGUIN-1 of said invention (1), and the rat gene which carries out the code of rat protein MAGUIN-2 of claim 2.

(4) The DNA fragment which is cDNA of the gene of said invention (3) and has the base sequence of the array number 2.

(5) The DNA fragment which is cDNA of the gene of said invention (3) and has the base sequence of the array number 4.

(6) The recombination vector which holds said invention (4) or the DNA fragment of (5).

(7) The antibody to rat protein MAGUIN-1 of said invention (1).

(8) The animal protein which has the amino acid sequence of the array number 1 and homologous, and has the same function as rat protein MAGUIN-1 of said invention (1).

(9) The animal protein which has the amino acid sequence of the array number 3 and homologous, and has

he same function as rat protein MAGUIN-2 of said invention (2).

0013] protein MAGUIN-1 of said invention (1) – yeast two-hybrid -- it is the new protein molecule isolated is a result of looking for an interaction molecule with S-SCAM using law, and it has a SAM domain, a PDZ domain, and PH domain, and has the consensus motif of association to PDZ in the C terminal, and PSD-95 are combined. Moreover, protein MAGUIN-2 of said invention (2) are alternative splicing form (isoform) of MAGUIN-1, and they are missing in the PDZ joint motif.

0014] It is checked as a result of the homology retrieval by the computer that such protein is new protein. In addition, it was named MAGUIN from "membrane-associated guanylate kinase-interacting protein".

0015] Hereafter, the gestalt of operation is explained in detail about said invention [of this application] (1) (9).

0016]

Embodiment of the Invention] Rat protein MAGUIN-1 of invention (1) is the protein by which the code was carried out to cDNA which showed the base sequence to the array number 2, and it has the amino acid sequence of the array number 1. And the SAM domain (amino acid number 8-75), the PDZ domain (amino acid number 157-296), PH domain (amino acid number 572-667), and the consensus motif (amino acid number 1029-1032) of association to PDZ are contained in this amino acid sequence.

0017] Rat protein MAGUIN-2 of invention (2) are the protein by which the code was carried out to cDNA which showed the base sequence to the array number 4, and they have the amino acid sequence of the array number 2. And the SAM domain (amino acid number 8-75), the PDZ domain (amino acid number 157-296), and PH domain (amino acid number 572-667) are included in this amino acid sequence.

0018] Moreover, as a result of NOZAN analysis, as shown in drawing 1 , having discovered MAGUIN-1 only in the brain of a rat is checked. These protein MAGUIN-1 and MAGUIN-2 (it may be hereafter indicated as "MAGUINs") are acquirable by the well-known approach, i.e., the method of isolating from the brain of a rat, the method of preparing a peptide by chemosynthesis based on the amino acid sequence offered by this invention, or the approach of producing by recombinant DNA technology using the cDNA fragment offered by this invention. For example, when acquiring MAGUINs by recombinant DNA technology, RNA is prepared by in vitro imprint from the vector which has invention (4) or the cDNA fragment of (5), and it can be discovered by in vitro one by performing an in vitro translation by making this into mold. Moreover, if a translation field is rearranged to a suitable expression vector by the well-known approach, cDNA can make MAGUINs which carries out a code discover in large quantities by Escherichia coli, the Bacillus subtilis, yeast, an animal cell, etc.

0019] In making the protein MAGUINs of this invention discover by microorganisms, such as Escherichia coli At least the origin which can be reproduced in a microorganism, a promotor, and a ribosome bond part To the expression vector which has a cDNA cloning part, a terminator, etc. If the obtained transformant is cultivated after creating the expression vector (Invention (6)) which carried out insertion association and rearranged the translation field of cDNA of this invention and carrying out the transformation of the host cell by this expression vector, cDNA can mass-produce within a microorganism MAGUINs which is carrying out the code. Or it can also be made discovered as a fusion protein with other protein. By cutting the obtained fusion protein by the suitable protease, cDNA can also acquire only the protein part which carries out a code.

0020] If the translation field of cDNA of this invention is rearranged to the expression vector for animal cells which has the promotor for animal cells, a splicing field, the (Pori A) addition part, etc. (invention (6)) and it introduces it in an animal cell in making Protein MAGUINs discover by the animal cell, MAGUINs of this invention can be discovered within an animal cell.

0021] The rat protein MAGUINs obtained by the approach as above can be used as an antigen for, for example, creating the antibody which recognizes this protein specifically.

0022] A peptide fragment (5 or more amino acid residue) including what kind of partial amino acid sequence of the amino acid sequence expressed with the array numbers 1 or 2 is also contained in the rat protein MAGUINs of this invention. These peptide fragments can also be used as an antigen for producing an antibody.

0023] the gene of the rat to which invention (3) carries out the code of the above-mentioned protein MAGUINs -- it is -- for example, invention (4), cDNA of (5), or its part -- it can isolate from the existing genomic library by using an array as a probe.

0024] Invention (4) and cDNA of (5) are characterized by having the base sequence expressed with the array numbers 2 and 4, respectively, and are carrying out the code of said protein MAGUIN-1 of this invention, and MAGUIN-2, respectively. Since Protein MAGUINs is discovered in the brain tissue of a rat at least, the same clone as cDNA of this invention can be easily obtained by screening a rat brain cDNA

ibrary using the oligonucleotide probe compounded based on the base sequence of the array numbers 2 or 4. or these oligonucleotides -- a primer -- carrying out -- polymerase chain reaction (PCR) -- the purpose cDNA is also compoundable using law.

[0025] Generally the polymorphism according [the gene of mammalian] to individual difference is accepted frequently. Therefore, in the array numbers 2 or 4, cDNA by which the permutation by 1 or addition of two or more nucleotides, deletion, and/or other nucleotides is made is also contained in this invention.

[0026] The protein with which similarly the permutation by 1 produced by these modification or addition of two or more amino acid residue, deletion, and/or other amino acid residue is made is also contained in this invention as long as it has the activity of the protein which has the amino acid sequence expressed with the array numbers 1 or 2.

[0027] Furthermore, a DNA fragment (10 or more bps) including what kind of partial base sequence of the base sequence expressed with the array numbers 2 and 4 or the DNA fragment which consists of those antisense strands is also contained in invention (4) and the DNA fragment of (5).

[0028] The antibody of invention (7) can be obtained as a polyclonal antibody or a monoclonal antibody by the well-known approach by using its protein itself or partial peptide as an antigen.

[0029] Invention (8) and the protein of (9) are animal proteins other than the rat which has the amino acid sequence of the array numbers 1 and 3 and homologous, and has invention (1) and the same function as MAGUINs of (2) respectively. Such protein can isolate cDNA of homologous for invention (4) or cDNA of (5) from the cDNA library of various animal species as a probe, and can acquire it by making this cDNA discover by the microorganism, the animal cell, or the plant cell.

[0030] Hereafter, the approach which isolated Protein MAGUINs, the analysis result of the function, etc. are explained in detail.

1. isolation of MAGUINs -- according to the well-known approach (J.Neurosci.16, 2488-2494, 1996), the rat brain yeast two-hybrid library (5x106 clones) was built, and it screened using the cutting tool (bait) including the PDZ domain of S-SCAM. Consequently, 34 positive clones were obtained, among those 22 pieces were new clones. Although two clones as which a manifestation is regarded only in a brain were furthermore obtained as a result of NOZAN analysis, it became clear that one clone was the rat homologue of a well-known human gene among those.

[0031] About the one remaining clones, all arrays were determined as screening of a rat brain cDNA library by PCR which uses the rat brain cDNA as mold. Two isoforms were obtained and it was respectively named MAGUIN-1 and MAGUIN-2. MAGUIN-1 is the protein by which the code was carried out to cDNA which showed the base sequence to the array number 2, and it has the amino acid sequence of the array number 1. And the SAM domain (amino acid number 8-75), the PDZ domain (amino acid number 157-296), PH domain (amino acid number 572-667), and the consensus motif (amino acid number 1029-1032) of association to PDZ are contained in this amino acid sequence. MAGUIN-2 are the protein by which the code was carried out to cDNA which showed the base sequence to the array number 4, and they have the amino acid sequence of the array number 2. And the SAM domain (amino acid number 8-75), the PDZ domain (amino acid number 157-296), and PH domain (amino acid number 572-667) are included in this amino acid sequence.

2. In order to investigate association with MAGUINs, S-SCAM and association place ** with PSD-95/SAP90, MAGUINs, and S-SCAM, an extract, and GST-MAGUIN -12 (C terminal of MAGUIN-1) or GST-MAGUIN -16 (C terminal of MAGUIN-2) of the COS cell which discovers S-SCAM was incubated. Although a result is as having been shown in drawing 2 and the C terminal of MAGUIN-1 combined with S-SCAM, the C terminal of MAGUIN-2 was not combined. In addition, in the Western blot analysis of drawing 2 , the anti-S-SCAM antibody (J.Biol.Chem.273, 21105-21110, 1998) was used for detection of a band.

[0032] In order to investigate the MAGUIN-1 joint field of S-SCAM, various Myc indicator S-SCAM constructs shown in drawing 3 A were created, and it incubated with GST-MAGUIN -12 (C terminal of MAGUIN-1). A result is as having been shown in drawing 3 B, and it was checked that the 4th and the 5th PDZ domain are related to both association. In addition, in the Western blot analysis of drawing 3 B, the anti-Myc monoclonal antibody nine E10 (from ATCC to acquisition) was used for detection of a band.

[0033] Next, PSD-95/SAP90 and each PDZ domain of PSD93/chapsyn110 and SAP97 were obtained by reverse yeast two-hybrid screening which makes a cutting tool pBTM116 MAGUIN-9 (846-1032 of the array number 1). From this, combining MAGUIN-1 not only with S-SCAM but with PSD-95/SAP90 and its isoform was checked. Moreover, drawing 4 A is as a result of the Western blot analysis which investigated ********** of MAGUIN-1, and PSD-95/SAP90. Triton X-100 extract of the rough synaptosome fraction of a rat was incubated on anti-PSD-95/SAP90 blood serum or the blood serum in front of immunity, and the Protein G-sepharose bead, and the immuno blot of the protein combined with the bead was carried out by the anti-

MAGUIN antibody. An anti-MAGUIN antibody is a rabbit polyclonal antibody created considering pGex4T-1 MAGUIN-1 (716-858 of the array number 1) as immunogen. In drawing 4 A, the sample to which the lane 1 incubated with the sample of the rough synaptosome extract before an incubation, and the lane 2 incubated with the blood serum in front of immunity, and a lane 3 are anti-PSD-95/SAP90 blood serum and the sample which incubated. As shown in this drawing 4 A, MAGUIN-1 **********(ed) with PSD-95 / SAP90 of he rough synaptosome fraction of a rat. Furthermore, drawing 4 B is as a result of the Western blot analysis which investigated association with MAGUIN-1 and the PDZ domain of PSD-95 / SAP90. The COS cell extract which discovers various PSD-95/SAP90 which carried out the Myc indicator, and GST-MAGUIN 12 (C terminal of MAGUIN-1) or GST-MAGUIN -16 (C terminal of MAGUIN-2) which combined the glutathione bead was incubated, and the protein combined with the bead was detected by the anti-Myc antibody. Lanes 1-3 are [the PDZ domain of PSD-95/SAP90 and the lane 7-9 of overall-length PSD-95/SAP90 and a lane 4-6] GK domains of the Src homology 3 and PSD-95/SAP90. Furthermore, lanes 1, 4, and 7 are [the sample after an incubation with GST-MAGUIN -12 and the lanes 3, 6, and 9 of the sample before an incubation and lanes 2, 5, and 8] the samples after an incubation with GST-MAGUIN -16. As shown in this drawing 4 B, it was checked that MAGUIN-1 combines with the PDZ domain of PSD-95 / SAP90.

3. As shown also in the organization and the cell distribution map 1 of MAGUINs, it was checked in NOZAN analysis that MAGUINs is discovered only in the brain of a rat. Furthermore, when the distribution kicked to a rat brain cell was investigated, mainly discovering MAGUINs by synapse plasmlemma (synaptic plasma membrane:SPM) and the PSD fraction was checked (drawing 5 A). Moreover, MAGUINs was distributed over the soma and the neural spine in the hippocampus neurone of a rat, and coexisting with NMDAR1 was checked by it (drawing 5 B).

[0034]
Effect of the Invention] The protein MAGUINs of this invention is new protein specifically combined with protein S-SCAM which has played the role important for the PSD formation in a nerve synaptic connection posterior part, and this protein is useful to the ED for the elucidation of mechanisms, such as generating of the nervous system of animals including Homo sapiens, development or storage, and study, a diagnosis of the various nervous diseases which consider the functional or structural failure of a nervous system as a cause further, prevention, and a therapy etc. as explained in detail above.

[0035]
Layout Table]
SEQUENCE LISTING <110> Japan Science and Technology Corporation <120> Protein MAGUIN <130> <140> <141> <160> 4<170> PatentIn Ver.2.0 <210> 1 <211> 1032<212> PRT <213> rat <400> 1 Met Ala Leu Ile Met Glu Pro Val Ser Lys Trp Ser Pro Ser Gln Val 1 5 10 15 Val Asp Trp Met Lys Gly Leu Asp Asp Cys Leu Gln Gln Tyr Ile Lys 20 25 30 Asn Phe Glu Arg Glu Lys Ile Ser Gly Asp Gln Leu Leu Arg Ile Thr 35 4045 His Gln Glu Leu Glu Asp Leu Gly Val Ser Arg Ile Gly His Gln Glu 50 55 60 Leu Ile LeuGlu Ala Val Asp Leu Leu Cys Ala Leu Asn Tyr Gly Leu 65 70 75 80 Glu ThrGlu Asn Leu Lys Thr Leu Ser His Lys Leu Asn Ala Ser Ala 85 90 95 Lys Asn Leu Gln Asn Phe Ile Thr Gly Arg Arg Arg Ser Gly His Tyr 100 105 110 Asp Gly Arg Thr Ser Arg Lys Leu Pro Asn Asp Phe Leu Thr Ser Val 115 120 125 Val Asp LeuIle Gly Ala Ala Lys Ser Leu Leu Ala Trp Leu Asp Arg 130 135 140 Ser Pro Phe Ala Ala Val Thr AspTyr Ser Val Thr Arg Asn Asn Val 145 150 155 160 Ile Gln Leu Cys Leu Glu Leu Thr Thr Ile Val Gln Gln Asp Cys Thr 165 170175 Val Tyr Glu Thr GluAsn Lys Ile Leu His Val Cys Lys Thr Leu Ser 180 185190 Gly Val Cys Asp His Ile Ile SerLeu Ser Ser Asp Pro Leu Val Ser 195 200 205 Gln Ser Ala His Leu Glu Val Ile Gln Leu Ala Asn Ile Lys Pro Ser 210 215 220 Glu Gly Leu Gly Met Tyr Ile Lys Ser Thr Tyr Asp Gly Leu His Val 225 230 235 240 Ile Thr Gly Thr Thr Glu Asn Ser Pro Ala Asp Arg Cys Lys Lys Ile 245 250 255His Ala Gly Asp Glu Val-Ile-Gln-Val-Asn His Gln Thr Val Val-Gly 260 265 270 Trp Gln Leu Lys Asn Leu Val Asn Ala-Leu-Arg-Glu-Asp-Pro-Ser-Gly 275 280 285 Val Ile Leu Thr Leu Lys Lys Arg Pro Gln Ser Met Leu Thr Ser Ala 290 295 300 Pro Ala Leu Leu Lys Asn Met Arg Trp Lys Pro Leu Ala Leu Gln Pro 305 310 315 320 Leu Ile Pro Arg Ser Pro Thr Ser Ser Val Ala Thr Pro Ser Ser Thr 325 330 335 Ile Ser Thr Pro Thr Lys Arg Asp Ser Ser Ala Leu Gln Asp Leu Tyr 340 345 350 Ile Pro Pro Pro Pro Ala Glu Pro Tyr Ile Pro Arg Asp Glu Lys Gly 355 360 365 Asn Leu Pro Cys Glu Asp Leu Arg Gly His Met Val Gly Lys Pro Val 370 375 380 His Lys Gly Ser Glu Ser Pro Asn Ser Phe Leu Asp Gln Glu Tyr Arg 385 390 395 400 Lys Arg Phe Asn Ile Val Glu Glu Asp Thr Val Leu Tyr Cys Tyr Glu 405 410 415 Tyr Glu Lys Gly Arg Ser Ser Ser Gln Gly Arg Arg Glu Ser Thr Pro 420 425 430 Thr Tyr Gly Lys Leu Arg Pro Ile Ser Met Pro Val Glu Tyr Asn Trp 435 440 445 Val Gly Asp Tyr Glu Asp Pro Asn Lys Met Lys Arg Asp Ser Arg Arg 450 455 460 Glu Asn Ser Leu Leu Arg Tyr Met Ser Asn Glu Lys Ile Ala Gln Glu 465 470 475 480 Glu Tyr Met Phe Gln Arg Asn Ser Lys Lys Asp Thr Gly Lys Lys Ser 485 490 495 Lys Lys Lys Gly Asp Lys Ser Thr Ser Pro Thr His Tyr Ser Leu Leu 500 505 510 Pro Ser Leu Gln Met Asp Ala

Leu Arg Gln Asp Ile Met Gly Thr Pro 515 520 525 Val Pro Glu Thr Thr Leu Tyr His Thr Phe Gln Gln Ser Ser
Leu Gln 530 535 540 His Lys Ser Lys Lys Lys Asn Lys Gly Ala Ile Ala Gly Lys Ser Lys 545 550 555 560 Arg
Arg Ile Ser Cys Lys Asp Leu Gly Arg Gly Asp Cys Glu Gly Trp 565 570 575 Leu Trp Lys Lys Lys Asp Ala
Lys Ser Tyr Phe Ser Gln Lys Trp Lys 580 585 590 Lys-Tyr-Trp-Phe-Val Leu Lys Asp Ala Ser-Leu-Tyr-Trp-
Tyr-Ile-Asn 595 600 605 Glu Glu Asp Glu Lys Ala Glu Gly Phe-Ile-Ser-Leu-Pro-Glu-Phe-Lys 610 615 620
Ile Asp Arg Ala Ser Glu Cys Arg Lys Lys Tyr Ala Phe Lys Ala Cys 625 630 635 640 His Pro Lys Ile Lys Ser
Phe Tyr Phe Ala Ala Glu His Leu Asp Asp 645 650 655 Met Asn Arg Trp Leu Asn Arg Ile Asn Met Leu Thr
Ala Gly Tyr Ala 660 665 670 Glu Arg Glu Arg Ile Lys Gln Glu Gln Asp Tyr Trp Ser Glu Ser Asp 675 680 685
Lys Glu Glu Ala Asp Thr Pro Ser Thr Pro Lys Gln Asp Ser Pro Pro 690 695 700 Pro Pro Tyr Asp Thr Tyr
Pro Arg Pro Pro Ser Met Ser Cys Ala Ser 705 710 715 720 Pro Tyr Val Glu Ala Lys His Ser Arg Leu Ser
Ser Thr Glu Thr Ser 725 730 735 Gln Ser Gln Ser Ser His Glu Glu Phe Arg Gln Glu Val Thr Gly Ser 740
745 750 Ser Ala Val Ser Pro Ile Arg Lys Thr Ala Ser Gln Arg Arg Ser Trp 755 760 765 Gln Asp Leu Ile Glu
Thr Pro Leu Thr Ser Ser Gly Leu His Tyr Leu 770 775 780 Gln Thr Leu Pro Leu Glu Asp Ser Val Phe Ser
Asp Ser Ala Ala Ile 785 790 795 800 Ser Pro Glu His Arg Arg Gln Ser Thr Leu Pro Thr Gln Lys Cys His 805
810 815 Leu Gln Asp His Tyr Gly Pro Tyr Pro Leu Ala Glu Ser Glu Arg Met 820 825 830 Gln Val Leu Asn
Gly Asn Gly Gly Lys Pro Arg Ser Phe Thr Leu Pro 835 840 845 Arg Asp Ser Gly Phe Asn His Cys Cys Leu
Asn Ala Pro Val Ser Ala 850 855 860 Cys Asp Pro Gln Asp Asp Ile Gln Pro Pro Glu Val Glu Glu Glu Glu
865 870 875 880 Glu Glu Glu Glu Glu Glu Ala Ala Gly Glu Asn Ile Gly Glu Lys Asn 885 890 895 Glu Asn
Arg Glu Glu Lys Leu Gly Asp Ser Leu Gln Asp Leu Tyr Arg 900 905 910 Ala Leu Glu Glu Ala Ser Leu Ser
Pro Leu Gly Glu His Arg Ile Ser 915 920 925 Thr Lys Ile Glu Tyr Lys Leu Ser Phe-Ile-Lys-Arg-Cys-Asn-Asp-
Pro 930 935 940 Val Met Asn Glu Lys Leu His Arg Leu-Arg-Ile-Leu-Lys-Ser Thr Leu 945 950 955 960 Lys A
la Arg Glu Gly Glu Val Ala Ile Ile Asp Lys Val Leu Asp Asn 965 970 975 Pro Asp Leu Thr Ser Lys Glu Phe
Gln Gln Trp Lys Gln Met Tyr Leu 980 985 990 Asp Leu Phe Leu Asp Ile Cys Gln Asn Thr Thr Ser Asn Asp
Pro Leu 995 1000 1005 Ser Ile Ser Ser Glu Val Asp Val Ile Thr Ser Ser Leu Thr His Thr 1010 1015 1020
His Ser Tyr IleGluThr His Val 1025 1030 <210> 2 <211> 3099 <212> DNA <213> rat <400> 2 atggctctga
taatggaacc agtgagcaaa tggtctccga gtcaagtagt ggactggatg 60 aaaggtctcg atgactgctt acagcagtat
attaagaactttgagaggga gaagatcagt 120 ggagaccagc tgctacgcat tacacatcag gagctagaag atctgggagt
cagccgtatt 180 ggccatcagg agctgatctt ggaggcagtc gaccttctgt gtgcactgaa ttatggctta 240 gaaacagaaa
atctaaaaac cctttctcac aagttgaatg cgtctgccaa aaatctacag 300 aacttcataa caggaaggag gaggagcggc cattatgatg
ggaggaccag ccgaaaatta 360 ccaaatgact ttttgacctc agttgtggat ctgattggag cagccaagag tcttcttgct 420
tggttggaca ggtcaccatt tgctgctgtg acagactatt cagtcacaag aaataatgtc 480 atacaactct gcttggagtt aacaacaatt
gtgcaacagg attgcactgt ctatgaaact 540 gaaaataaaa ttcttcatgt gtgtaaaact ctttctgggg tctgtgacca catcatatcc 600
ctgtcgtcag atcctctggt ttcacagtct gctcaccttg aagtgattca gctggcaaat 660 atcaaaccaa gtgaagggct gggtatgtat
attaaatcca catatgatgg gctccatgta 720 attacgggaa ccacagaaaa ttcacctgca gatcggtgca agaaaatcca
cgctggggat 780 gaagtgattc aagtcaatca tcagactgtg gtggggtggc agttgaaaaa tttggtgaat 840 gcactacgag
aggacccgag tggtgttatc ttaactttga aaaagcgacc tcagagcatg 900 cttacctcag caccagcttt actgaaaaat atgagatgga
agccccttgc tctgcagcct 960 cttataccta gaagtcccac aagcagcgtt gccacgcctt ccagcaccat cagtacaccc
1020accaaaagag acagttctgc tctccaggat ctctacatcc cccctcctcc tgcagaacca 1080tacattccca gggacgaaaa
aggaaacctt ccttgtgaag atctcagagg acatatggtg 1140ggcaagccag ttcataaggg atctgaatcc cccaattcat ttctggacca
ggaatatcga 1200 aagaggttta acattgttga agaagacact gtcttatatt gctatgaata tgaaaaagga 1260cgatcaagta g
tcaaggaag-acgagaaagc acaccaacct-atggcaagct acgacctata 1320tctatgccag tcgaatataa-ttgggtgggg
gactatgaag-atccaaataa gatgaagaga 1380gatagtagaa gagaaaactc-tctacttcga tacatgagca-atgaaaaaat
tgctcaagaa 1440gaatacatgt ttcagagaaa-cagcaaaaag gacacaggga-agaagtcaaa aaagaagggt 1500gataagagta
ctagcccaac tcactactcc ttgctgccta gtttgcaaat ggatgcattg 1560agacaagaca tcatgggcac accagtgcca
gaaaccacac tgtaccatac atttcagcag 1620tcctctctgc agcacaaatc aaagaagaaa aacaaggag ctatagcagg
caagagcaaa 1680agaagaattt cttgcaagga tcttggccgt ggtgactgtg aaggatggct ttggaaaaag 1740aaagatgcaa
agagctattt ttcacagaaa tggaagaagt actggtttgt cctaaaggat 1800gcatccttgt actggtatat taatgaagag gatgaaaaag
cagagggatt catcagtctg 1860cctgaattta aaatcgatag agccagtgaa tgccgcaaga aatatgcatt caaggcctgc
1920catcctaaaa tcaaaagctt ttatttcgct gctgagcatc tcgatgacat gaacaggtgg 1980cttaatagaa ttaacatgct
gactgcagga tatgccgaaa gagagaggat taaacaagaa 2040caagattact ggagtgagag tgacaaagaa gaagcagata
ctccatcaac accaaaacaa 2100gacagccctc cacccccata tgacacatac ccacggcctc cctctatgag ttgtgccagt
2160ccttatgtgg aagcaaaaca gagccggctt tcctccacag agacctctca gtctcagtct 2220tcacatgagg aatttcgaca
ggaagttact gggagcagtg ctgtgtcccc cattcgcaag 2280acagccagtc agcgccgctc ctggcaggat ttaatcgaga
cgccactgac aagttcaggc 2340ttacactatc ttcagactct gcctctggaa gattctgtct ctctgactc ggcagccatc
2400tcccctgagc acaggcggca gtccactctt ccaactcaga agtgccacct ccaggatcat 2460tatggcccat accctttagc
tgagagcgag aggatgcaag tgttgaatgg caatgggggc 2520aagcctcgaa gtttactct gcctcgagat agcgggttca
accattgctg tctgaacgca 2580ccagttagtg cctgcgatcc acaagatgac atacagccac cagaggtgga ggaagaagag

2640gaggaggagg aagaggaggc agcaggagaa aacataggcg agaaaaatga aaatagggaa 2700gaaaagttag jagactcctt gcaagattta tacagggcac tggaggaagc cagtctgtca 2760ccactaggag aacatcgcat ttcaaccaag atcgaataca aactatcgtt cataaaacga 2820tgcaatgatc ctgtgatgaa tgagaagctg cacaggctga gaattcttaa aagcacttta 2880aaggccagag aaggggaagt agccattatt gataaagtcc tagacaatcc agacttgaca 2940tctaaagaat tccaacaatg jaagcagatg tacctcgacc ttttcttgga tatctgtcaa 3000aacaccacct caaatgaccc actaagtatt tcttctgaag tagatgttat cacttcctct 3060 ctaacacata ctcattcgta cattgagaca cacgtgtaa 3099 <210> 3 <211> 896<212> PRT <213> rat <400> 3 Met Ala Leu Ile Met Glu Pro Val Ser Lys Trp Ser-Pro-Ser-Gln-Val 1 5 10 15 Val Asp Trp Met Lys-Gly-Leu-Asp-Asp Cys Leu Gln Gln Tyr-Ile-Lys 20 25 30 Asn Phe Glu Arg Glu Lys Ile Ser Gly Asp Gln Leu Leu Arg Ile Thr 35 40 45 His Gln Glu Leu Glu Asp Leu Gly Val Ser Arg Ile Gly His Gln Glu 50 55 60 Leu Ile Leu Glu Ala Val Asp Leu Leu Cys Ala Leu Asn Tyr Gly Leu 65 70 75 80 Glu Thr Glu AsnLeu Lys Thr Leu Ser His Lys Leu Asn Ala Ser Ala 85 90 95 Lys Asn Leu GlnAsn Phe Ile Thr Gly Arg Arg Arg Ser Gly His Tyr 100 105 110 Asp Gly Arg ThrSer Arg Lys Leu Pro Asn Asp Phe Leu Thr Ser Val 115 120 125 Val Asp Leu Ile Gly Ala Ala Lys Ser Leu Leu Ala Trp Leu Asp Arg 130 135 140 Ser Pro Phe Ala Ala Val Thr Asp Tyr Ser Val Thr Arg Asn Asn Val 145 150 155 160 Ile Gln Leu Cys Leu Glu Leu Thr Thr Ile Val Gln Gln Asp Cys Thr 165 170 175 Val Tyr Glu ThrGlu Asn Lys Ile Leu His Val Cys Lys Thr Leu Ser 180 185 190 Gly Val Cys AspHisIle Ile Ser Leu Ser Ser Asp Pro Leu Val Ser 195 200 205 Gln Ser Ala His Leu Glu Val Ile Gln Leu Ala Asn Ile Lys Pro Ser 210 215 220 Glu Gly Leu Gly Met Tyr Ile Lys Ser Thr Tyr Asp Gly Leu His Val 225 230 235 240 Ile Thr Gly Thr Thr Glu Asn Ser Pro Ala Asp Arg Cys Lys Lys Ile 245 250 255 His Ala Gly AspGluVal Ile Gln Val Asn His Gln Thr Val Val Gly 260 265 270 Trp Gln Leu Lys Asn Leu ValAsnAla Leu Arg Glu Asp Pro Ser Gly 275 280 285 Val Ile Leu Thr Leu Lys Lys Arg Pro Gln Ser Met Leu Thr Ser Ala 290 295 300 Pro Ala Leu Leu Lys Asn Met Arg Trp Lys Pro Leu Ala Leu Gln Pro 305 310 315 320 Leu Ile Pro Arg Ser Pro Thr-Ser-Ser-Val-Ala Thr Pro Ser Ser Thr 325 330 335 Ile Ser Thr Pro Thr Lys Arg Asp Ser-Ser-Ala-Leu-Gln-Asp-Leu-Tyr 340 345 350 Ile Pro Pro Pro Pro Ala Glu Pro Tyr Ile Pro Arg Asp Glu Lys Gly 355 360365 Asn Leu Pro Cys Glu Asp Leu Arg Gly His Met Val Gly Lys Pro Val 370 375 380 His Lys Gly Ser Glu Ser Pro Asn Ser Phe Leu Asp Gln Glu Tyr Arg 385 390 395 400 Lys Arg Phe Asn Ile Val Glu Glu Asp Thr Val Leu Tyr Cys Tyr Glu 405 410 415 Tyr Glu Lys Gly Arg Ser Ser Ser Gln Gly Arg Arg Glu Ser Thr Pro 420 425 430 Thr Tyr Gly Lys Leu Arg Pro Ile Ser Met Pro Val Glu Tyr Asn Trp 435 440 445 Val Gly Asp Tyr Glu Asp Pro Asn Lys Met Lys Arg Asp Ser Arg Arg 450 455 460 Glu Asn Ser Leu Leu Arg Tyr Met Ser Asn Glu Lys Ile Ala Gln Glu 465 470 475 480 Glu Tyr Met Phe Gln Arg Asn Ser Lys Lys Asp Thr Gly Lys Lys Ser 485 490 495 Lys Lys Lys Gly Asp Lys Ser Thr Ser Pro Thr His Tyr Ser Leu Leu 500 505 510 Pro Ser Leu Gln Met Asp Ala Leu Arg Gln Asp Ile Met Gly Thr Pro 515 520 525 Val Pro Glu Thr Thr Leu Tyr His Thr Phe Gln Gln Ser Ser Leu Gln 530 535 540 His Lys Ser Lys Lys Lys Asn Lys Gly Ala Ile Ala Gly Lys Ser Lys 545 550 555 560 Arg Arg Ile Ser Cys Lys Asp Leu Gly Arg Gly Asp Cys Glu Gly Trp 565 570 575 Leu Trp Lys Lys Asp Ala Lys Ser Tyr Phe Ser Gln Lys Trp Lys 580 585 590 Lys Tyr Trp Phe Val Leu Lys Asp Ala Ser Leu Tyr Trp Tyr Ile Asn 595 600 605 Glu Glu Asp Glu Lys Ala Glu Gly Phe Ile Ser Leu Pro Glu Phe Lys 610 615 620 Ile Asp Arg Ala Ser Glu Cys Arg Lys Lys Tyr Ala Phe Lys Ala Cys 625 630 635 640 His Pro Lys Ile Lys Ser Phe Tyr Phe Ala Ala Glu His Leu Asp Asp 645 650 655 Met Asn-Arg-Trp-Leu Asn Arg Ile Asn Met-Leu-Thr-Ala-Gly-Tyr-Ala 660 665 670 Glu Arg Glu Arg Ile Lys Gln Glu Gln-Asp-Tyr-Trp-Ser-Glu-Ser-Asp 675 680 685 Lys Glu Glu A la Asp Thr Pro Ser Thr Pro Lys Gln Asp Ser Pro Pro 690 695 700 Pro Pro Tyr Asp Thr Tyr Pro Arg Pro Pro Ser Met Ser Cys Ala Ser 705 710 715 720 Pro Tyr Val Glu Ala Lys His Ser Arg Leu Ser Ser Thr Glu Thr Ser 725 730 735 Gln Ser Gln Ser Ser His Glu Glu Phe Arg Gln Glu Val Thr Gly Ser 740 745 750 Ser Ala Val Ser Pro Ile Arg Lys Thr Ala Ser Gln Arg Arg Ser Trp 755 760 765 Gln Asp Leu Ile Glu Thr Pro Leu Thr Ser Ser Gly Leu His Tyr Leu 770 775 780 Gln Thr Leu Pro Leu Glu Asp Ser Val Phe Ser Asp Ser Ala Ala Ile 785 790 795 800 Ser Pro Glu His Arg Arg Gln Ser Thr Leu Pro Thr Gln Lys Cys His 805 810 815 Leu Gln Asp His Tyr Gly Pro Tyr Pro Leu Ala Glu Ser Glu Arg Met 820 825 830 Gln Val Leu Asn Gly Asn Gly Gly Lys Pro Arg Ser Phe Thr Leu Pro 835 840 845 Arg Asp Ser Gly Phe Asn His Cys Cys Leu Asn Ala Pro Val Ser Ala 850 855 860 Cys Asp Pro Gln Asp Asp Ile Gln Pro Pro Glu 865 870 875 880 Glu Glu Glu Glu Glu Glu Ala Ala Gly Glu Asn Ile Gly Glu Lys Ser 885 890 895 <210> 4<211> 4462 <212> DNA <213> rat <220> <221> CDS <223>.. (463) (3153) <400> 4 ggtccgcagg aagccctcgt catttccgcc gggcggctcg tcgtcgtgcc gggggctcca 60 ctcccgcgcg ccaggcgagc gagcgggcaa gttggctgcg gccgtgcgcc ggcgtctctt 120 cccttccccg ccgcgaccct cggcggctcg agccatgaac tgaagcccct cgggacgcac 180 accccgggct gagccgcgga cgctgcagct gcggtggcag cctgggctgg gacagagccg 240 accccggcgc gttatcc cgg-ccggccggct-ggccgtcgcc gctcccaccg-cggccgcctc 300 cgacgctcca gcctccggac ggacacacaa-aagccaccct ctccctgtcg-cccacccccag 360 gaggctgcgg ccggccactg accccacctg-agggcgcagc ggggctgcgga-cttggctttc 420 tgtccgagct ccgcgtccca cgcggcacca-accccgccgc ccatggctct-gataatggaa 480 ccagtgagca aatggtctcc gagtcaagta gtggactgga tgaaaggtct cgatgactgc 540 ttacagcagt atattaagaa ctttgagagg gagaagatca gtggagacca gctgctacgc 600 attacacatc aggagctaga agatctggga gtcagccgta ttggccatca

ggagctgatc 660 ttggaggcag tcgaccttct gtgtgcactg aattatggct tagaaacaga aaatctaaaa 720 accctttctc
acaagttgaa tgcgtctgcc aaaaatctac agaacttcat aacaggaagg 780 aggaggagcg gccattatga tgggaggacc
agccgaaaat taccaaatga cttttttgacc 840 tcagttgtgg atctgattgg agcagccaag agtcttcttg cttggttgga caggtcacca
900 tttgctgctg tgacagacta ttcagtcaca agaaataatg tcatacaact ctgcttggag 960 ttaacaacaa ttgtgcaaca
ggattgcact gtctatgaaa ctgaaaataa aattcttcat 1020 gtgtgtaaaa ctctttctgg ggtctgtgac cacatcatat ccctgtcgtc
agatcctctg 1080 gtttcacagt ctgctcacct tgaagtgatt cagctggcaa atatcaaacc aagtgaaggg 1140 ctgggtatgt
atattaaatc cacatatgat gggctccatg taattacggg aaccacagaa 1200 aattcacctg cagatcggtg caagaaaatc
cacgctgggg atgaagtgat tcaagtcaat 1260 catcagactg tggtgggggtg gcagttgaaa aatttggtga atgcactacg
agaggacccg 1320 agtggtgtta tcttaacttt gaaaaagcga cctcagagca tgcttacctc agcaccagct 1380 ttactgaaaa
atatgagatg gaagcccctt gctctgcagc ctcttatacc tagaagtccc 1440 acaagcagcg ttgccacgcc ttccagcacc
atcagtacac ccaccaaaag agacagttct 1500 gctctccagg atctctacat cccccctcct cctgcagaac catacattcc
cagggacgaa 1560 aaaggaaacc ttccttgtga agatctcaga ggacatatgg tgggcaagcc agttcataag 1620 ggatctgaat
cccccaattc attctggac caggaatatc gaaagaggtt taacattgtt 1680 gaagaagaca ctgtcttata ttgctatgaa tatgaaaaag
gacgatcaag tagtcaagga 1740 agacgagaaa gcacaccaac ctatggcaag ctacgaccta tatctatgcc agtcgaatat 1800
aattgggtgg gggactatga agatccaaat aagatgaaga gagatagtag aagagaaaac 1860 tctctacttc gatacatgag
caatgaaaaa attgctcaag aagaatacat gtttcagaga 1920 aacagcaaaa aggacacagg gaagaagtca aaaaagaagg
gtgataagag tactagccca 1980 actcactact ccttgctgcc tagtttgcaa atggatgcat tgagacaaga catcatgggc 2040
acaccagtgc cagaaaccac actgtaccat acatttcagc agtcctctct gcagcacaaa 2100 tcaaagaaga aaaacaaagg
agctatagca ggcaagagca aaagaagaat ttcttgcaag 2160 gatcttggcc gtggtgactg tgaaggatgg ctttggaaaa
agaaagatgc aaagagctat 2220 ttttcacaga aatggaagaa gtactggttt gtcctaaagg atgcatcctt gtactggtat 2280
attaatgaag aggatgaaaa agcagaggga ttcatcagtc tgcctgaatt taaaatcgat 2340 agagccagtg aatgccgcaa
gaaatatgca ttcaaggcct gccatcctaa aatcaaaagc 2400 ttttatttcg ctgctgagca tctcgatgac atgaaca ggt-ggcttaatag-
aattaacatg 2460ctgactgcag gatatgccga aagagagagg-attaaacaag aacaagatta-ctggagtgag 2520agtgacaaag
aagaagcaga tactccatca-acaccaaaac aagacagccc-tccacccca 2580tatgacacat acccacggcc tccctctatg-
agttgtgcca gtccttatgt-ggaagcaaaa 2640cacagccggc tttcctccac agagacctct cagtctcagt cttcacatga ggaatttcga
2700 caggaagtta ctgggagcag tgctgtgtcc cccattcgca agacagccag tcagcgccgc 2760 tcctggcagg atttaatcga
gacgccactg acaagttcag gcttacacta tcttcagact 2820 ctgcctctgg aagattctgt cttctctgac tcggcagcca tctcccctga
gcacaggcgg 2880 cagtccactc ttccaactca gaagtgccac ctccaggatc attatggccc taccctttta 2940 gctgagagcg
agaggatgca agtgttgaat ggcaatgggg gcaagcctcg aagttttact 3000 ctgcctcgag atagcgggtt caaccattgc
tgtctgaacg caccagttag tgcctgcgat 3060 ccacaagatg acatacagcc accagaggtg gaggaagaag aggaggagga
ggaagaggag 3120 gcagcaggag aaaacatagg cgagaaaagc taatacactg cgagagttgg tagaacctct 3180 ccatgccata
tcggatccac ttctgttggc actcaacccca ttggaatcac agattgataa 3240 gctaatgttt agagaattta gatggaagag agtcggcacg
gcgcagactc aacatcaacc 3300 tcttgcaagc aactaaaatg gcctcgtcct tgctgtttat aacagaaaac agacttgtaa 3360
aaagcttaga tcatcaagtg ttgtggattt ggggcctcta aagggatatt aagaggggca 3420 ggccactctt aagaagaatt cgagctttca
acaatgggac tagcataaga tcaaagccaa 3480 tcaagatgga ccacagtagc cgagaactgc cctttttatg ggagaacaga
gggtaaaggg 3540 tcttgcctgg ggaagggttc tgtggcgaca cctcagttgt cttctcctga catcaggaag 3600 agagggctca
gcaccaatga ttagaaaatc ctggctgtta aatggactct tggtgggatg 3660 gcaagactct tcattgttaa ggcaaagcag
cggaagcaaa ttctgtgtta agtgtatttg 3720 cattttttcaa atttgacatg ctgtattgta ctaaattaag tgtaatctat taaggcaagg 3780
tgtacacacc ctgctcagaa actcaatatg tttattctat tattaaagca tattcaggtg 3840 taaggcagac tgctttcagt gacactaatg
aaggaaattt ctcatgccag gcattattag 3900 attcctcagc gaaaaatcct aactttctcc ttttccttgg cacttacata gaagtgatgt
3960 tgcttcataa gctggaatga gctattcttt ttctgtaaaa tattccaatc tataggctgt 4020 ggttttccat atttgaaaaa gtattttgtc
tggatgtctt tcaaattagc ttcagatatt 4080 atttcatact ctgtaactgg gcccccgtgg ctcaatattg cttctttttt cttctgtggt 4140
agatgtgaaa ttttcctttag ttggatgagg tacactgtaa tgattttaat gctaattaat 4200 gaatatttca tacttgtgca gtgaaaaaat
aatttaatat tgtattatga aatgttttc 4260 ttcctgttgc tgcagtgtgt ggtattgcat aatatgaata cctgtaaaaa atataaattg 4320
cttaaaaata aaaatattac cagttggtac cagatctttt tagatagcta aataaatttgc 4380 taagtatgct tgatagcaaa tatctttcta
acaaagaaaa atctatgctt ctttctcaca 4440 aatgaaatct cgtgccgaat tc 4462

[Translation done.]